

Data Preparation for Case 31

Complete Version

Jae Jung

2025-01-05

Libraries

Loading common packages

```
# install.packages("haven")
library(tidyverse)
library(haven)
library(labelled)
library(sjlabelled)
library(janitor)
library(sjPlot)
library(psych)
set_theme(theme_minimal())
```

Setting Up Data

Reading SPSS data

```
data <- read_sav("Case 31 Data_efcu_1.sav")
data <- data |>
  clean_names()

# generate data dictionary
data |>
  view_df()
```

Table 1: Data frame: data

ID	Name	Label	Values	Value Labels
1	q3	Employees are Courteous	1	Strongly
			5	Disagree
2	q4	employees are helpful	1	Strongly
			5	Disagree
3	q5	employees are professional	1	Strongly
			5	Disagree
4	q6	employees are available	1	Strongly
			5	Disagree
5	q7	savings rates	1	Strongly Agree
			2	Very High
			3	High
			4	Average
			5	Low
			6	Very Low
6	q8	statement frequency	1	No Opinion
			2	Very High
			3	High
			4	Average
			5	Low
			6	Very Low
7	q9	statement frequency	1	No Opinion
			2	Too Often
			3	Very Often
			4	About Right
			5	Not Often
8	q10	statement accuracy	1	Enough
			2	Never
			3	Excellent
			4	Good
9	q11	statement understandable	1	Fair
			2	Poor
10	q12	Account Confidential	1	Yes
			2	No

11	q13	regular share account	1 2	Aware and Have Used Aware but Have Not Used
12	q14	special subaccounts	1 2	Aware and Have Used Aware but Have Not Used
13	q15	Christmas club account	1 2	Aware and Have Used Aware but Have Not Used
14	q16	IRA	1 2	Aware and Have Used Aware but Have Not Used
15	q17	Master credit card	1 2	Aware and Have Used Aware but Have Not Used
16	q18	signature loans	1 2	Aware and Have Used Aware but Have Not Used
17	q19	new car loans	1 2	Aware and Have Used Aware but Have Not Used
18	q20	late model car loans	1 2	Aware and Have Used Aware but Have Not Used
19	q21	older model car loans	1 2	Aware and Have Used Aware but Have Not Used
20	q22	household goods loans	1 2	Aware and Have Used Aware but Have Not Used

21	q23	recreational loans	1 2	Aware and Have Used Aware but Have Not Used
22	q24	share collateralized loan	1 2	Aware and Have Used Aware but Have Not Used
23	q25	IRA loans	1 2	Aware and Have Used Aware but Have Not Used
24	q26	line of credit loans	1 2	Aware and Have Used Aware but Have Not Used
25	q27	currently have loan	1 2	Yes No
26	q30	CU rates lower	1 5	Strongly Disagree Strongly Agree
27	q31	CU confidential	1 5	Strongly Disagree Strongly Agree
28	q32	CU is prompt	1 5	Strongly Disagree Strongly Agree
29	q33	CU meets needs of members	1 5	Strongly Disagree Strongly Agree
30	q34	CU loan applications simple	1 5	Strongly Disagree Strongly Agree
31	q37	overall rating	1 2 3 4 5 6	Excellent Good Average Poor Very Poor No Opinion
32	overall_employee_performace		<i>range: 2.8-5.0</i>	
33	q38	headquarter	1 2	yes no

Data Preparation

Attitudes towards Employee Proficiency

q3

```
data |> # ctr + shift + M
  count(q3) |>
  mutate(percent = n/sum(n))
```

```
# A tibble: 3 x 3
  q3          n percent
<dbl+lbl>   <int>   <dbl>
1 3          13  0.102
2 4          57  0.445
3 5 [Strongly Agree] 58  0.453
```

```
five_pt_likert <- c("Strongly Disagree" = 1,
                    "Disagree" = 2,
                    "Uncertain" = 3,
                    "Agree" = 4,
                    "Strongly Agree" = 5)

data_clean <- data |>
  set_value_labels(q3 = five_pt_likert) |>
  set_variable_labels(q3 = "Employees are Courteous")

data_clean |>
  select(q3) |>
  var_label()
```

```
$q3
[1] "Employees are Courteous"
```

```
data_clean |>
  select(q3) |>
  val_labels() |>
  unique()
```

```
# A tibble: 3 x 1
  q3
  <dbl+lbl>
1 4 [Agree]
2 5 [Strongly Agree]
3 3 [Uncertain]
```

q4

```
data_clean <- data_clean |>
  set_variable_labels(q4 = "Employees are helpful") |>
  set_value_labels(q4 = five_pt_likert)

data_clean |>
  select(q4) |>
  var_label()
```

```
$q4
[1] "Employees are helpful"
```

```
data_clean |>
  select(q4) |>
  val_labels()
```

```
# A tibble: 128 x 1
  q4
  <dbl+lbl>
1 4 [Agree]
2 4 [Agree]
3 4 [Agree]
4 4 [Agree]
```

```

5 4 [Agree]
6 4 [Agree]
7 5 [Strongly Agree]
8 5 [Strongly Agree]
9 5 [Strongly Agree]
10 4 [Agree]
# i 118 more rows

```

q5 and q6

```

data_clean <- data_clean |>
  set_variable_labels(q5 = "Employees are professional") |>
  set_value_labels(q5 = five_pt_likert) |>
  set_variable_labels(q6 = "Employees are available") |>
  set_value_labels(q6 = five_pt_likert)

data_clean |>
  select(q5, q6) |>
  var_label()

```

```

$q5
[1] "Employees are professional"

```

```

$q6
[1] "Employees are available"

```

```

data_clean |>
  select(q5, q6) |>
  val_labels()

```

```

# A tibble: 128 x 2
   q5          q6
  <dbl+lbl> <dbl+lbl>
1 4 [Agree]    4 [Agree]
2 4 [Agree]    4 [Agree]
3 4 [Agree]    4 [Agree]
4 3 [Uncertain] 4 [Agree]

```

```

5 4 [Agree]          4 [Agree]
6 4 [Agree]          4 [Agree]
7 5 [Strongly Agree] 5 [Strongly Agree]
8 4 [Agree]          4 [Agree]
9 5 [Strongly Agree] 5 [Strongly Agree]
10 4 [Agree]         4 [Agree]
# i 118 more rows

```

Reliability analysis

```

data_clean |>
  select(q3:q6) |>
  alpha() #.87: cutoff is .70

```

Reliability analysis

Call: alpha(x = select(data_clean, q3:q6))

raw_alpha	std.alpha	G6(smc)	average_r	S/N	ase	mean	sd	median_r
0.87	0.87	0.85	0.62	6.6	0.019	4.2	0.59	0.65

95% confidence boundaries

	lower	alpha	upper
Feldt	0.82	0.87	0.9
Duhachek	0.83	0.87	0.9

Reliability if an item is dropped:

	raw_alpha	std.alpha	G6(smc)	average_r	S/N	alpha	se	var.r	med.r
q3	0.85	0.86	0.80	0.66	5.9	0.022	0.00056	0.67	
q4	0.79	0.79	0.73	0.56	3.8	0.033	0.00800	0.57	
q5	0.82	0.83	0.80	0.62	4.8	0.028	0.01894	0.67	
q6	0.85	0.85	0.80	0.66	5.8	0.023	0.00610	0.68	

Item statistics

	n	raw.r	std.r	r.cor	r.drop	mean	sd
q3	128	0.80	0.81	0.73	0.66	4.4	0.66
q4	128	0.91	0.91	0.88	0.82	4.3	0.70
q5	128	0.85	0.85	0.78	0.73	4.1	0.66

q6 128 0.83 0.82 0.73 0.67 4.1 0.79

Non missing response frequency for each item

	2	3	4	5	miss
q3	0.00	0.10	0.45	0.45	0
q4	0.01	0.11	0.45	0.44	0
q5	0.00	0.18	0.56	0.26	0
q6	0.03	0.18	0.47	0.32	0

Factor analysis

Factor Analysis

```
data_clean |>
  select(q3:q6) |>
  factanal(factors = 1, rotation = "none")
```

Call:

```
factanal(x = select(data_clean, q3:q6), factors = 1, rotation = "none")
```

Uniquenesses:

	q3	q4	q5	q6
	0.431	0.141	0.423	0.473

Loadings:

	Factor1
q3	0.754
q4	0.927
q5	0.760
q6	0.726

	Factor1
SS loadings	2.531
Proportion Var	0.633

Test of the hypothesis that 1 factor is sufficient.

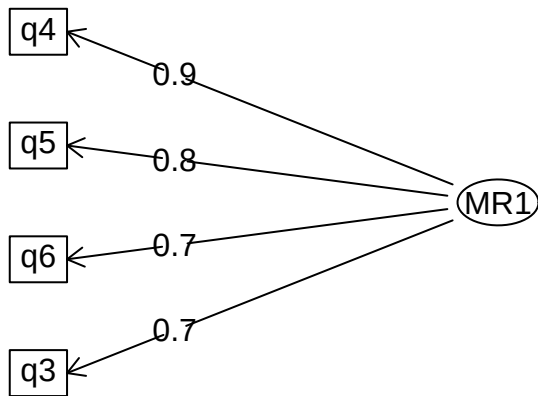
The chi square statistic is 11.35 on 2 degrees of freedom.

The p-value is 0.00344

Diagram

```
data_clean |>
  select(q3:q6) |>
  fa(nfactors = 1, rotate = "none") |>
  fa.diagram()
```

Factor Analysis



PCA (Principle Component Analysis)

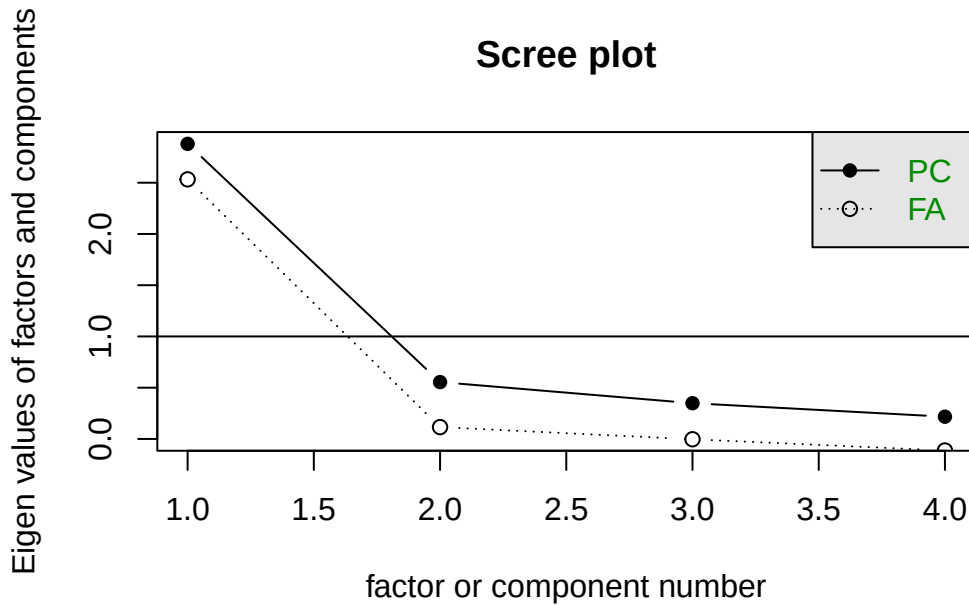
```
data_clean |>
  select(q3:q6) |>
  prcomp(scale. = TRUE) |>
  summary() # % of variance explained by the first factor
```

Importance of components:

	PC1	PC2	PC3	PC4
Standard deviation	1.6968	0.7450	0.5906	0.46601
Proportion of Variance	0.7198	0.1388	0.0872	0.05429
Cumulative Proportion	0.7198	0.8585	0.9457	1.00000

Scree plot

```
data_clean |>
  select(q3:q6) |>
  scree()
```



Creating a composite scale

```
data_clean <- data_clean |>
  mutate(at_emp = rowMeans(across(q3:q6), na.rm = TRUE))

# from psych: descriptive statistics
describe(data_clean$at_emp)
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
X1	1	128	4.21	0.59	4	4.25	0.74	2.75	5	2.25	-0.32	-0.69	0.05

q7r: Savings Rate

```
data_clean |> count(q7)
```

```
# A tibble: 5 x 2
  q7      n
  <dbl+lbl> <int>
1 2 [High]    22
2 3 [Average] 83
3 4 [Low]     12
4 5 [Very Low] 1
5 NA         10
```

```
data_clean <- data_clean |>
  mutate(q7r = case_when(q7 == 1 ~ 5,
                        q7 == 2 ~ 4,
                        q7 == 3 ~ 3,
                        q7 == 4 ~ 2,
                        q7 == 5 ~ 1,
                        TRUE ~ NA)) |>
  set_variable_labels(q7r = "Savings rates") |>
  set_value_labels(q7r = c("Very Low" = 1,
                          "Low" = 2,
                          "Average" = 3,
                          "High" = 4,
                          "Very High" = 5))

data_clean |> count(q7r)
```

```
# A tibble: 5 x 2
  q7r      n
  <dbl+lbl> <int>
1 1 [Very Low]    1
2 2 [Low]        12
3 3 [Average]     83
4 4 [High]       22
5 NA           10
```

q8r: Fund borrowing rates

```
data_clean |> count(q8)
```

```
# A tibble: 6 x 2
  q8      n
  <dbl> <int>
1 1 [Very High]    4
2 2 [High]        26
3 3 [Average]      59
4 4 [Low]         17
5 5 [Very Low]     1
6 NA              21
```

```
data_clean <- data_clean |>
  mutate(q8r = case_when(q8 == 1 ~ 5,
                        q8 == 2 ~ 4,
                        q8 == 3 ~ 3,
                        q8 == 4 ~ 2,
                        q8 == 5 ~ 1,
                        TRUE ~ NA)) |>
  set_variable_labels(q8r = "Fund borrowing rates") |>
  set_value_labels(q8r = c("Very Low" = 1,
                          "Low" = 2,
                          "Average" = 3,
                          "High" = 4,
                          "Very High" = 5))

data_clean |> count(q8r)
```

```
# A tibble: 6 x 2
  q8r      n
  <dbl> <int>
1 1 [Very Low]    1
2 2 [Low]        17
3 3 [Average]     59
4 4 [High]       26
5 5 [Very High]   4
6 NA            21
```

q9r: Statement Frequency

```
data_clean |> count(q9)
```

```
# A tibble: 2 x 2
  q9          n
  <dbl+lbl> <int>
1 3 [About Right]    99
2 4 [Not Often Enough] 29
```

```
data_clean |>
  select(q9) |>
  class()
```

```
[1] "tbl_df"      "tbl"        "data.frame"
```

```
data_clean <- data_clean |>
  mutate(q9r = 6 - q9) |>
  set_variable_labels(q9r = "Statement Frequency") |>
  set_value_labels(q9r = c("Never" = 1,
                           "Not Often Enough" = 2,
                           "About Right" = 3,
                           "Very Often" = 4,
                           "Too Often" = 5))
data_clean |> count(q9r)
```

```
# A tibble: 2 x 2
  q9r          n
  <dbl+lbl> <int>
1 2 [Not Often Enough]    29
2 3 [About Right]        99
```

q10r: Statement Accuracy

```
data_clean |> count(q10)
```

```
# A tibble: 4 x 2
  q10      n
  <dbl> <int>
1 1 [Excellent] 68
2 2 [Good]      54
3 3 [Fair]       4
4 4 [Poor]       2
```

```
data_clean <- data_clean |>
  mutate(q10r = 5 - q10) |>
  set_variable_labels(q10r = "Statement Accuracy") |>
  set_value_labels(q10r = c("Poor" = 1,
                             "Fair" = 2,
                             "Good" = 3,
                             "Excellent" = 4))
data_clean |> count(q10r)
```

```
# A tibble: 4 x 2
  q10r      n
  <dbl> <int>
1 1 [Poor]      2
2 2 [Fair]      4
3 3 [Good]     54
4 4 [Excellent] 68
```

Awareness of Financial Services

```
aware_scale_3 <- c("Unaware" = 1,
                   "Aware but Have Not Used" = 2,
                   "Aware and Have Used" = 3)
```

q13r

```
data_clean |> count(q13)
```

```
# A tibble: 3 x 2
  q13      n
  <dbl+lbl> <int>
1 1 [Aware and Have Used]    13
2 2 [Aware but Have Not Used] 32
3 3                          83
```

```
data_clean <- data_clean |>
  mutate(q13r = 4 - q13) |>
  set_variable_labels(q13r = "Regular share account") |>
  set_value_labels(q13r = aware_scale_3)

data_clean |> count(q13r)
```

```
# A tibble: 3 x 2
  q13r      n
  <dbl+lbl> <int>
1 1 [Unaware]    83
2 2 [Aware but Have Not Used] 32
3 3 [Aware and Have Used]    13
```

Creating a function

```
library(rlang)
```

Attaching package: 'rlang'

The following objects are masked from 'package:sjlabelled':

as_character, as_label

The following objects are masked from 'package:purrr':

```
%%, flatten, flatten_chr, flatten_dbl, flatten_int, flatten_lgl,  
flatten_raw, invoke, splice
```

```
aware_recode <- function(tbl, var, var_label) {  
  
  varr <- paste0(rlang::as_string(ensym(var)), "r")  
  
  tbl <- tbl |>  
  mutate(!!varr := 4 - as.numeric({{var}})) |>  
  set_variable_labels(!!varr := var_label) |>  
  set_value_labels(!!varr := aware_scale_3)  
  
  return(tbl)  
}
```

```
data_clean <- data_clean |>  
  aware_recode(q15, "Special subaccounts")
```

```
data_clean |>  
  count(q15r)
```

```
# A tibble: 3 x 2  
  q15r          n  
  <dbl+lbl>    <int>  
1 1 [Unaware]    42  
2 2 [Aware but Have Not Used] 75  
3 3 [Aware and Have Used]    11
```

q14r

```
data_clean |> count(q14)
```

```
# A tibble: 3 x 2  
  q14          n  
  <dbl+lbl>    <int>  
1 1 [Aware and Have Used]    12
```

2	2	[Aware but Have Not Used]	70
3	3		46

```
data_clean <- data_clean |>
  aware_recode(q14, "Special subaccounts")

data_clean |> count(q14r)
```

```
# A tibble: 3 x 2
  q14r          n
  <dbl+lbl>    <int>
1 1 [Unaware]    46
2 2 [Aware but Have Not Used] 70
3 3 [Aware and Have Used]    12
```

q15r-q26

functions that didn't work

q15r

```
data_clean |> count(q15)
```

```
# A tibble: 3 x 2
  q15          n
  <dbl+lbl>    <int>
1 1 [Aware and Have Used]    11
2 2 [Aware but Have Not Used] 75
3 3          42
```

```
data_clean <- data_clean |>
  aware_recode(q15, "Christmas club account")

data_clean |> count(q15r)
```

```
# A tibble: 3 x 2
  q15r      n
  <dbl+lbl> <int>
1 1 [Unaware]    42
2 2 [Aware but Have Not Used] 75
3 3 [Aware and Have Used]    11
```

q16r

```
data_clean |> count(q16)
```

```
# A tibble: 3 x 2
  q16      n
  <dbl+lbl> <int>
1 1 [Aware and Have Used]    29
2 2 [Aware but Have Not Used] 64
3 3                        35
```

```
data_clean <- data_clean |>
  aware_recode(q16, "IRA")
data_clean |> count(q16r)
```

```
# A tibble: 3 x 2
  q16r      n
  <dbl+lbl> <int>
1 1 [Unaware]    35
2 2 [Aware but Have Not Used] 64
3 3 [Aware and Have Used]    29
```

q17r

```
data_clean |> count(q17)
```

```
# A tibble: 3 x 2
```

	q17	n
	<dbl+lbl>	<int>
1	1 [Aware and Have Used]	56
2	2 [Aware but Have Not Used]	57
3	3	15

```
data_clean <- data_clean |>
  aware_recode(q17, "Master credit card")

data_clean |> count(q17r)
```

```
# A tibble: 3 x 2
```

	q17r	n
	<dbl+lbl>	<int>
1	1 [Unaware]	15
2	2 [Aware but Have Not Used]	57
3	3 [Aware and Have Used]	56

q18r

```
data_clean |> count(q18)
```

```
# A tibble: 3 x 2
```

	q18	n
	<dbl+lbl>	<int>
1	1 [Aware and Have Used]	19
2	2 [Aware but Have Not Used]	104
3	3	5

```
data_clean <- data_clean |>
  aware_recode(q18, "Signature loans")

data_clean |> count(q18r)
```

```
# A tibble: 3 x 2
  q18r                                n
  <dbl+lbl>                        <int>
1 1 [Unaware]                        5
2 2 [Aware but Have Not Used]      104
3 3 [Aware and Have Used]          19
```

q19r

```
data_clean |> count(q19)
```

```
# A tibble: 3 x 2
  q19                                n
  <dbl+lbl>                        <int>
1 1 [Aware and Have Used]          24
2 2 [Aware but Have Not Used]      99
3 3                                5
```

```
data_clean <- data_clean |>
  aware_recode(q19, "New car loans")

data_clean |> count(q19r)
```

```
# A tibble: 3 x 2
  q19r                                n
  <dbl+lbl>                        <int>
1 1 [Unaware]                        5
2 2 [Aware but Have Not Used]      99
3 3 [Aware and Have Used]          24
```

q20r

```
data_clean |> count(q20)
```

```
# A tibble: 3 x 2
  q20          n
  <dbl+lbl>    <int>
1 1 [Aware and Have Used]    14
2 2 [Aware but Have Not Used] 102
3 3                        12
```

```
data_clean <- data_clean |>
  aware_recode(q20, "Late model car loans")

data_clean |> count(q20r)
```

```
# A tibble: 3 x 2
  q20r          n
  <dbl+lbl>    <int>
1 1 [Unaware]    12
2 2 [Aware but Have Not Used] 102
3 3 [Aware and Have Used]    14
```

q21r

```
data_clean |> count(q21)
```

```
# A tibble: 3 x 2
  q21          n
  <dbl+lbl>    <int>
1 1 [Aware and Have Used]    12
2 2 [Aware but Have Not Used] 83
3 3                        33
```

```
data_clean <- data_clean |>
  aware_recode(q21, "Older model car loans")

data_clean |> count(q21r)
```

```
# A tibble: 3 x 2
  q21r          n
  <dbl+lbl>    <int>
1 1 [Unaware]    33
2 2 [Aware but Have Not Used] 83
3 3 [Aware and Have Used]    12
```

q22r

```
data_clean |> count(q22)
```

```
# A tibble: 3 x 2
  q22          n
  <dbl+lbl>    <int>
1 1 [Aware and Have Used]    11
2 2 [Aware but Have Not Used] 80
3 3                          37
```

```
data_clean <- data_clean |>
  aware_recode(q22, "Household goods loans")
```

```
data_clean |> count(q22r)
```

```
# A tibble: 3 x 2
  q22r          n
  <dbl+lbl>    <int>
1 1 [Unaware]    37
2 2 [Aware but Have Not Used] 80
3 3 [Aware and Have Used]    11
```

q23r

```
data_clean |> count(q23)
```

```
# A tibble: 3 x 2
  q23          n
  <dbl+lbl>    <int>
```

1	1	[Aware and Have Used]	13
2	2	[Aware but Have Not Used]	60
3	3		55

```
data_clean <- data_clean |>
  aware_recode(q23, "Recreational loans")

data_clean |> count(q23r)
```

```
# A tibble: 3 x 2
  q23r          n
  <dbl+lbl>    <int>
1 1 [Unaware]    55
2 2 [Aware but Have Not Used] 60
3 3 [Aware and Have Used]    13
```

q24r

```
data_clean |> count(q24)
```

```
# A tibble: 3 x 2
  q24          n
  <dbl+lbl>    <int>
1 1 [Aware and Have Used]    2
2 2 [Aware but Have Not Used] 57
3 3                69
```

```
data_clean <- data_clean |>
  aware_recode(q24, "Share collateralized loan")

data_clean |> count(q24r)
```

```
# A tibble: 3 x 2
  q24r          n
  <dbl+lbl>    <int>
1 1 [Unaware]    69
2 2 [Aware but Have Not Used] 57
3 3 [Aware and Have Used]    2
```

q25r

```
data_clean |> count(q25)
```

```
# A tibble: 3 x 2
  q25          n
  <dbl+lbl>    <int>
1 1 [Aware and Have Used]      4
2 2 [Aware but Have Not Used]  46
3 3                          78
```

```
data_clean <- data_clean |>
  aware_recode(q25, "IRA loans")
```

```
data_clean |> count(q25r)
```

```
# A tibble: 3 x 2
  q25r          n
  <dbl+lbl>    <int>
1 1 [Unaware]      78
2 2 [Aware but Have Not Used]  46
3 3 [Aware and Have Used]      4
```

q26r

```
data_clean |> count(q26)
```

```
# A tibble: 2 x 2
  q26          n
  <dbl+lbl>    <int>
1 1 [Aware and Have Used]      48
2 2 [Aware but Have Not Used]  80
```

```
data_clean <- data_clean |>
  aware_recode(q26, "Line of credit loans")
```

```
data_clean |> count(q26r)
```

```
# A tibble: 2 x 2
  q26r          n
  <dbl+lbl>    <int>
1 2 [Aware but Have Not Used]    80
2 3 [Aware and Have Used]       48
```

q30 - q34

```
data_clean <- data_clean |>
  set_value_labels(q30 = five_pt_likert,
                  q31 = five_pt_likert,
                  q32 = five_pt_likert,
                  q33 = five_pt_likert,
                  q34 = five_pt_likert)
```

q37r

```
data_clean |> count(q37)
```

```
# A tibble: 6 x 2
  q37          n
  <dbl+lbl>    <int>
1 1 [Excellent]    31
2 2 [Good]         71
3 3 [Average]      16
4 4 [Poor]         2
5 5 [Very Poor]    1
6 NA              7
```

```
data_clean <- data_clean |>
  mutate(q37r = case_when(q37 == 1 ~ 5,
                          q37 == 2 ~ 4,
                          q37 == 3 ~ 3,
                          q37 == 4 ~ 2,
                          q37 == 5 ~ 1,
```

```

      q37 == 6 ~ NA)) |>
set_value_labels(q37r = c("Very Poor" = 1,
      "Poor" = 2,
      "Average" = 3,
      "Good" = 4,
      "Excellent" = 5))

data_clean |> count(q37r)

```

```

# A tibble: 6 x 2
  q37r      n
  <dbl+lbl> <int>
1 1 [Very Poor] 1
2 2 [Poor] 2
3 3 [Average] 16
4 4 [Good] 71
5 5 [Excellent] 31
6 NA 7

```

Renaming Clean Data

```

data_preped <- data_clean

data_preped |>
  view_df()

```

Table 2: Data frame: data_preped

ID	Name	Label	Values	Value Labels
1	q3	Employees are Courteous	1	Strongly
			2	Disagree
			3	Disagree
			4	Uncertain
			5	Agree
				Strongly Agree

2	q4	Employees are helpful	1	Strongly
			2	Disagree
			3	Disagree
			4	Uncertain
			5	Agree
3	q5	Employees are professional		Strongly Agree
			1	Strongly
			2	Disagree
			3	Disagree
			4	Uncertain
4	q6	Employees are available	5	Agree
				Strongly Agree
			1	Strongly
			2	Disagree
			3	Disagree
5	q7	savings rates	4	Uncertain
			5	Agree
				Strongly Agree
			1	Very High
			2	High
6	q8	statement frequency	3	Average
			4	Low
			5	Very Low
			6	No Opinion
			1	Very High
7	q9	statement frequency	2	High
			3	Average
			4	Low
			5	Very Low
			6	No Opinion
8	q10	statement accuracy	1	Too Often
			2	Very Often
			3	About Right
			4	Not Often
			5	Enough
9	q11	statement understandable		Never
			1	Excellent
			2	Good
			3	Fair
			4	Poor
			1	Yes
			2	No

10	q12	Account	1	Yes
		Confidential	2	No
11	q13	regular share	1	Aware and Have
		account	2	Used
				Aware but Have
				Not Used
12	q14	special	1	Aware and Have
		subaccounts	2	Used
				Aware but Have
				Not Used
13	q15	Christmas club	1	Aware and Have
		account	2	Used
				Aware but Have
				Not Used
14	q16	IRA	1	Aware and Have
			2	Used
				Aware but Have
				Not Used
15	q17	Master credit	1	Aware and Have
		card	2	Used
				Aware but Have
				Not Used
16	q18	signature loans	1	Aware and Have
			2	Used
				Aware but Have
				Not Used
17	q19	new car loans	1	Aware and Have
			2	Used
				Aware but Have
				Not Used
18	q20	late model car	1	Aware and Have
		loans	2	Used
				Aware but Have
				Not Used
19	q21	older model car	1	Aware and Have
		loans	2	Used
				Aware but Have
				Not Used
20	q22	household goods	1	Aware and Have
		loans	2	Used
				Aware but Have
				Not Used

21	q23	recreational loans	1	Aware and Have
			2	Used
22	q24	share collateralized loan	1	Aware but Have
			2	Not Used
23	q25	IRA loans	1	Aware and Have
			2	Used
24	q26	line of credit loans	1	Aware but Have
			2	Not Used
25	q27	currently have loan	1	Aware and Have
			2	Used
26	q30	CU rates lower	1	Aware but Have
			2	Not Used
			3	Yes
			4	No
			5	Strongly
27	q31	CU confidential	1	Disagree
			2	Disagree
			3	Uncertain
			4	Agree
			5	Strongly Agree
28	q32	CU is prompt	1	Strongly
			2	Disagree
			3	Disagree
			4	Uncertain
			5	Agree
29	q33	CU meets needs of members	1	Strongly Agree
			2	Strongly
			3	Disagree
			4	Disagree
			5	Uncertain
				Agree
				Strongly Agree

30	q34	CU loan applications simple	1	Strongly
			2	Disagree
			3	Disagree
			4	Uncertain
			5	Agree
31	q37	overall rating		Strongly Agree
			1	Excellent
			2	Good
			3	Average
			4	Poor
			5	Very Poor
32	q38	overall_employee_performace headquarter	6	No Opinion
33			range: 2.8-5.0	
34	employeeperformance	at_emp	1	yes
			2	no
35			range: 2.8-5.0	
36	q7r	Savings rates	range: 2.8-5.0	
			1	Very Low
			2	Low
			3	Average
			4	High
37	q8r	Fund borrowing rates	5	Very High
			1	Very Low
			2	Low
			3	Average
			4	High
38	q9r	Statement Frequency	5	Very High
			1	Never
			2	Not Often
			3	Enough
			4	About Right
39	q10r	Statement Accuracy	5	Very Often
				Too Often
			1	Poor
			2	Fair
40	q13r	Regular share account	3	Good
			4	Excellent
			1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used

41	q15r	Christmas club account	1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used
42	q14r	Special subaccounts	1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used
43	q16r	IRA	1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used
44	q17r	Master credit card	1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used
45	q18r	Signature loans	1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used
46	q19r	New car loans	1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used
47	q20r	Late model car loans	1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used
48	q21r	Older model car loans	1	Unaware
			2	Aware but Have
			3	Not Used
				Aware and Have
				Used

49	q22r	Household goods loans	1	Unaware
			2	Aware but Have
			3	Not Used
50	q23r	Recreational loans	1	Aware and Have
			2	Used
			3	Unaware
51	q24r	Share collateralized loan	1	Aware but Have
			2	Not Used
			3	Aware and Have
52	q25r	IRA loans	1	Used
			2	Unaware
			3	Aware but Have
53	q26r	Line of credit loans	1	Not Used
			2	Aware and Have
			3	Used
54	q37r		1	Very Poor
			2	Poor
			3	Average
			4	Good
			5	Excellent

Exporting Prepped Data

SPSS data

```
data_preped |>
  write_sav("Case 31 Data_efcu_Prepped.sav")
```

Excel Data as Metadata

```
library(writexl)

data_preped |>
  write_xlsx("Case 31 Data_efcu_Prepped.xlsx")
```

Appendix

alternative way using `select()`: in this case, we need to use anonymous function

```
data_clean |>
  (\(df) df |>
    mutate(row_mean = rowMeans(select(df, q3, q4, q5, q6), na.rm = TRUE))
  )() |>
  select(row_mean)
```

Alternative way to create a recode function for awareness

```
aware_recode <- function(tbl, var, var_label) {

  varr <- paste0(deparse(substitute(var)), "r")

  tbl <- tbl |>
  mutate(!!sym(varr) := 4 - {{var}}
    ) |>
  set_variable_labels(!!sym(varr) := var_label) |>
  set_value_labels(!!sym(varr) := aware_scale_3)

  return(tbl)
}
```